TGCat#: THE CHANDRA TRANSMISSION GRATING DATA CATALOG AND ARCHIVE

Citation: Huenemoerder, David P., Arik Mitschang, Daniel Dewey, Michael A. Nowak, Norbert S. Schulz, Joy S. Nichols, John E. Davis, et al. "TGCat#: THE CHANDRA TRANSMISSION GRATING DATA CATALOG AND ARCHIVE." The Astronomical Journal 141, no. 4 (March 8, 2011): 129. © 2011 The American Astronomical Society

As Published: http://dx.doi.org/10.1088/0004-6256/141/4/129

Publisher: IOP Publishing

Persistent URL: http://hdl.handle.net/1721.1/93140

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

The Astronomical Journal, 141:129 (10pp), 2011 April doi:10.1088/0004-6256/141/4/129

*TGCat**: THE *CHANDRA* TRANSMISSION GRATING DATA CATALOG AND ARCHIVE

David P. Huenemoerder[1], Arik Mitschang[2], Daniel Dewey[1], Michael A. Nowak[1], Norbert S. Schulz[1], Joy S. Nichols[2], John E. Davis[1], John C. Houck[1], Herman L. Marshall[1], Michael S. Noble[1], Doug Morgan[2], and Claude R. Canizares[1]
[1] Massachusetts Institute of Technology, Kavli Institute for Astrophysics and Space Research, 70 Vassar St., Cambridge, MA 02139, USA
[2] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA

Received 2011 January 9; accepted 2011 February 11; published 2011 March 8

ABSTRACT

The *Chandra* Transmission Grating Data Archive and Catalog (*TGCat*) provides easy access to analysis-ready products, specifically, high-resolution X-ray count spectra and their corresponding calibrations. The web interface makes it easy to find observations of a particular object, type of object, or type of observation; to quickly assess the quality and potential usefulness of the spectra from pre-computed summary plots; or to customize a view with an interactive plotter, optionally combining spectra over multiple orders or observations. Data and responses can be downloaded as a package or as individual files, and the query results themselves can be retrieved as ASCII or Virtual Observatory tables. Portable reprocessing scripts used to create the archive and which use the *Chandra* X-ray Center's (CXC's) software and other publicly available software are also available, facilitating standard or customized reprocessing from Level 1 CXC archival data to spectra and responses with minimal user interaction.

Key words: catalogs – techniques: spectroscopic – X-rays: general

Online-only material: machine-readable and VO tables

1. INTRODUCTION

The *Chandra X-ray Observatory* (*CXO*) has two high-resolution grating instruments, the High Energy Transmission Grating Spectrometer (HETGS; Canizares et al. 2005) and the Low Energy Transmission Grating Spectrometer (LETGS; Brinkman et al. 2000). Given the excellent spatial resolution of the *Chandra* mirror assembly, these spectrometers also provide commensurately high spectral resolution. Weisskopf et al. (2002) give an overview of the observatory and some early scientific results. Paerels & Kahn (2003) have reviewed the scientific importance and impact of high-resolution X-ray spectroscopy in particular as enabled by the *Chandra* grating spectrometers (as well as by the complementary reflection grating spectrometers on *XMM-Newton*). This era represents the beginning of routine high-resolution X-ray spectroscopy for astronomy.

We found several compelling motivations for developing a catalog and archive of *Chandra* grating data. There have been several independent efforts that sought to go beyond the basic information and data provided by the *Chandra* archive. There are "HotGAS," a "Chandra Grating Spectroscopy Database for Active Galactic Nuclei,"[3] and "X-Atlas," a "Chandra Spectral Atlas" (Westbrook et al. 2008).[4] These are mostly interpretive catalogs which display galleries and derived characteristics for specific classes of astrophysical sources. For *XMM-Newton* high-resolution X-ray spectra, we have the "Browsing Interface for RGS Data" (BiRD) which provides simple search and spectral plotting of observations. These were all considered in our decision to develop *TGCat*. Our particular reasons are as follows.

1. We intend to provide a legacy of calibrated and analysis-ready high-resolution X-ray spectral products. *TGCat* provides the binned counts spectrum in a "Pulse Height Analyzer" (PHA) spectrum file format (both Types I and II), the effective area ("Auxiliary Response File" or ARF), and the line-spread function ("Response Matrix File" or RMF)[5] for each extraction, and with these one can perform quantitative spectral analysis and modeling. The primary *Chandra* X-ray Center (CXC) archive does not include instrumental responses (ARF and RMF) because those may be extraction-dependent, or multiple extractions from a single field. *TGCat* includes responses for each extraction as well as multiple extractions within a field.
2. We wish to make all of *Chandra*'s high-resolution X-ray spectra highly visible for browsing, via pre-computed summary plots and with easy and flexible web-based interactive plotting and output. In order to facilitate astrophysically motivated queries based on broad research goals, *TGCat* has a catalog interface by class of source, with fairly general and flexible search criteria.
3. The catalog and archive will make high-resolution X-ray spectra easily accessible. Provision of analysis-ready products removes a potential barrier of detailed instrumental knowledge and substantial software expertise required to extract spectra and to compute associated responses. Provision of summary plots and tables with a flexible choice of flux units encourages use of *Chandra* grating spectra in multi-band investigations, perhaps for which the high-energy region is not the researcher's primary area of expertise, or for which detailed X-ray spectroscopic modeling is not required.
4. A high-resolution X-ray spectral archive will help us prepare for future missions. The currently envisioned *International X-ray Observatory* (*IXO*) is fundamentally a high-resolution X-ray spectroscopy mission. Hence, it is important to facilitate and encourage use of high-resolution X-ray spectra within the high-energy astrophysics community. High-resolution spectra can be (or can be perceived to be) more difficult to analyze than CCD-resolution spectra.

---

* http://tgcat.mit.edu/

[3] http://hotgas.pha.jhu.edu/

[4] http://cxc.harvard.edu/XATLAS/

[5] For details of these FITS file format specifications, see http://heasarc.gsfc.nasa.gov/docs/heasarc/ofwg/ofwg_recomm.html.

Table 1
Observational Input Files

Type	Description
`asol1`	Aspect solution
`bpix1`	Bad pixel list
`dtf1`	Dead time factors (HRC only)
`evt1`	Event list
`flt1`	Good time intervals (GTI)
`msk1`	Detector active region
`pbk0`	Exposure parameters (ACIS only)
`stat1`	Exposure statistics (ACIS only)

Note. The "Type" is a *Chandra* Data Archive content string with a data-processing level digit appended.

However, the information content and diagnostic power of HETGS or LETGS spectra (resolving powers of up to 1000) can be much higher than CCD-resolution spectra (resolving power of about 50). Given the relatively long lead time to any *IXO* launch, it is important to grow and maintain a vibrant spectroscopic community.

5. *Chandra* grating observations represent relatively few total objects (∼1000), so there is reasonable ability to customize spectral extractions. Human input is invoked to review every extraction.
6. We can provide systematic and standardized processing parameters and products. With some exceptions for special cases (as recorded in a processing report), the database contains uniform and consistent extractions in terms of calibration and user-selectable processing parameters.

The catalog and archive can be accessed at the URL http://tgcat.mit.edu/. Here, we will describe the production, management, and public interface to the data. Some details on the system design have been presented previously by Huenemoerder et al. (2009) and by Mitschang et al. (2010).

2. ARCHIVE PRODUCTION AND ADMINISTRATION

The primary source for *TGCat* input data is the *Chandra* Data Archive (CDA).[6] The CDA is a robust, secure, permanent, and validated repository. We retrieve the minimum set of files from the CDA required to reprocess from "Level 1" (event lists and supporting data; see Table 1) into binned spectra, responses, and summary graphical and tabular products ("Level 2"). Each *Chandra* observation is labeled with an observation identifier (ObsID). Within each observation, an extraction is generated for each source of interest, as defined by a celestial location and corresponding spatial and spectral extraction regions.

The *Chandra* high-resolution spectrometers are objective grating designs and the focal-plane detectors are two-dimensional photon-counting arrays. Hence, the basic observational data are event lists giving time, position, and energy. The detector arrays are large enough to collect both positive and negative orders as well as the zeroth-order image. (Any other sources in the field will also be similarly diffracted.) The spectral extraction entails time-dependent coordinate transformations, spatial and energy filtering, and binning into a one-dimensional counts histogram which is the primary product for detailed spectral analysis. The radiometric calibration (equivalent to information obtained from flux standard star observations in ground-based optical spectroscopy) is contained in the effective area file ("ARF"), based on ground and flight calibration. The line-spread function (equivalent to the information from an emission-line lamp observation with a ground-based spectrometer) is contained in the RMF, also determined by ground and flight calibrations. Both of these have observation-dependent components, and both are required for detailed spectral analysis and modeling. (For more details on the spectrometers, consult the "Proposers' Observatory Guide."[7])

Initially, our sources correspond to the observer's proposed target, which is typically a well known, distinct, and relatively bright object. There are exceptions, however, such as for well separated multiple sources in a cluster or unrelated objects found in the observed field, bright knots of spatially complex extended sources, or potentially even a coordinate with no visible X-ray source extracted for determination of flux upper limits or for stacking over many exposures of the same field (or the same kind of object). In Table 1, we list and define the input products we retrieve from the CDA on a per-ObsID basis.

The retrieval, processing, and review are performed through automatic processes which monitor the CDA for publicly released data, spawn retrieval, initiate processing, request manual review by qualified scientists, and then automatically ingest into the public *TGCat* archive.

2.1. Processing Details

The event filtering, event transformation, spectral extraction, and response generation are done with standard CIAO programs (Fruscione et al. 2006) in conjunction with the *Chandra* Calibration Database[8] (CALDB). The methods are equivalent to those outlined in the CXC data analysis "Science Threads."[9] All of the processing steps can also be reproduced by users with distributed CXC software. A top-level pipeline sequence is shown in Table 2 and the output products are listed in Table 3. Most of the files are standard FITS format products whose definitions can be found on CIAO web pages (e.g., event, PHA, response files, aspect histograms, and light curves). Others added by the *TGCat* reprocessing scripts to augment the standard CIAO tool products are text summary files, summary plots, and a spectral properties file.

The processing is managed by a suite of custom scripts written in ISIS, the "Interactive System for Interpretation of Spectra"[10] (Houck & Denicola 2000), which determine instrumental and observational parameters from the data, configure the CIAO commands with defaults appropriate to grating spectral extraction or with observationally dependent parameters, and spawn the CIAO processes. ISIS uses the S-Lang scripting language[11] as its interpreter. S-Lang is an excellent choice for scientific analysis, being small, efficient, modular, and extensible (see Noble & Nowak 2008). For *TGCat* work with *Chandra* data, the `cfitsio`[12] module is particularly useful in providing a function-oriented interface to FITS-format (Hanisch et al. 2001) files. The primary advantage for using ISIS as the platform instead of the S-Lang shell is that ISIS provides much infrastructure for working with spectra and responses; it also includes a plotting package through the `pgplot`[13] module. We use the spectral support of ISIS to compute fluxes, plot spectra, and generate flux properties tables. In addition, the fitting support allows us to

---

[6] http://cda.harvard.edu/chaser/

[7] http://cxc.harvard.edu/proposer/POG/html/index.html

[8] http://cxc.harvard.edu/caldb/

[9] http://cxc.harvard.edu/ciao/threads/index.html

[10] http://space.mit.edu/cxc/isis/

[11] http://www.jedsoft.org/slang/

[12] http://heasarc.gsfc.nasa.gov/docs/software/fitsio/fitsio.html

[13] http://www.astro.caltech.edu/∼tjp/pgplot/

Table 2
Pipeline Processing Steps

Process[a]	Description
setup	Retrieve files from archive; configure directory. Set any non-standard extraction parameters (e.g., for serendipitous, crowded, or off-axis sources).
`*_process_events`	Detector-dependent; "`*`" is either `acis` or `hrc`.
`dmcopy`	Filter out bad events.
"detect" source	Either `tgdetect`, `findzo`, or "dead reckoning."
`tg_create_mask`	Define spectral regions in sky coordinates.
`tg_resolve_events`	Compute wavelengths and orders for each event.
`tgextract`	Spatially filter and bin the spectra by grating type, order, source, and background.
extract light curve	Uses `dmextract` for HRC, or `aglc` for ACIS; does source and background regions.
`asphist`	Compute aspect histograms (used for time-integrated effective area computation).
`mkgrmf`	Make grating Response Matrices (RMFs; valid for point sources, on-axis).
`mkgarf`	Make grating Auxiliary Response Files (ARFs, or effective area table).
`dmarfadd`	Combine the per-chip response files into one ARF per grating, order.
make summary plots	For internal verification and validation, and for catalog browsing.
make flux properties table	Counts and fluxes in bands (no modeling).

Notes. [a] The process is given in italics if it is a schematic name for several steps, or if there are several specific program options. Otherwise, a program name is given.

Table 3
Output Files

File Name	Description
`evt0.par`	Temporary observation file, output of `dmmakepar` (ASCII)
`evt1`	Filtered output of `acis_process_events` and `dmcopy` filters (FITS)
`evt1 _0`	Temporary `evt1`—output of `acis_process_events` (FITS)
`evt1 _1`	Temporary `evt1`—output of `dmcopy` (FITS)
`evt1 _2`	Temporary `evt1`—output of `dmcopy` (FITS)
`evt2`	Output of `tg_resolve_events` (FITS)
`findzo.ps`	(Optional) output of `findzo` (if ACIS) (postscript)
`lc`	Light curve, output of `aglc` (if ACIS), or `dmextract` (if LETG/HRC) (FITS)
`lc_bg`	Light curve, background (if LETG/HRC) (FITS)
`obs_config.txt`	Basic configuration info, from a header (ASCII)
`pha2`	Binned spectra; output of `tgextract` (FITS Type II format))
`tg_n.pha`	Binned spectra (FITS Type I format; split from `pha2` files).
`pha2 _bg`	Binned background spectrum (if LETG/HRC); both orders (FITS)
`pha2 _bg_-1`	Binned background spectrum for LETG/HRC, order −1 (FITS)
`pha2 _bg_1`	Binned background spectrum for LETG/HRC, order +1 (FITS)
`reg1a`	Sky region mask; output of `tg_create_mask` (FITS)
`sk.asphist`[a]	Aspect histograms, from `asphist` (FITS)
`src1a`	Source table; output of `tgdetect` (FITS)
`summary.tbl`	Table giving some of the observational information and some rates (ASCII)
`summary_fprops.fits`	Spectral flux properties table; counts and rates in bands (FITS)
`summary_tg_*.fits`[b]	Order sorting images (ACIS only, intermediate files; FITS)
`summary_*.ps`	Summary graphical products, spectra, and images (postscript)
`tg_n.arf`[b]	Grating ARFs (FITS)
`tg_n.pha`[b]	Type I PHA files (split from `pha2`, for convenience) (FITS)
`tg_n.rmf`[b]	Grating RMF files (FITS)

Notes.
[a] *sk*: means the detector subsystem chip, as in `s3`.
[b] *tg*: is a grating type, one of `heg`, `meg`, or `leg`, and *n* is the grating order, a signed integer.

run "findzo"(Nichols et al. 2008), a script which determines the zeroth-order centroid (the origin of the wavelength scale) to high accuracy even if the zeroth-order image is distorted by detector or telemetry saturation, or even excluded from telemetry (about 20% of the observations require the use of `findzo`). ISIS was designed especially for *Chandra* grating spectroscopic analysis and modeling and is actively used and supported for users by CXC grating scientists. Further details regarding the design and

Table 4
Flux Properties Table Example

Label	wmid (Å)	wlo (Å)	whi (Å)	count_rate (counts s^{-1})	photon_flux (photons cm^{-2} s^{-1})	energy_flux (erg cm^{-2} s^{-1})	Flag
heg_band	8.35	1.70	15.0	8.32E-02	1.41E-03	2.99E-12	0
meg_band	13.3	1.70	25.0	1.00E-01	3.35E-03	4.97E-12	0
letgs_band	81.0	2.00	160.	1.00E-01	3.91E-03	5.27E-12	1
letg_acis_band	26.0	2.00	50.0	1.00E-01	3.91E-03	5.27E-12	1
c1750	1.75	1.70	1.80	4.19E-05	2.91E-06	3.24E-14	0
Fe25	1.85	1.80	1.90	1.04E-04	5.21E-06	5.56E-14	0
FeK	1.95	1.90	2.00	2.09E-05	7.49E-07	7.53E-15	0
c2500	2.50	2.00	3.00	1.48E-03	2.56E-05	2.05E-13	0
...							
Ne10	12.1	12.1	12.2	5.91E-03	1.15E-04	1.89E-13	0
c13200	13.2	13.0	13.4	1.67E-03	4.62E-05	6.96E-14	0
...							
csc_b	13.4	2.00	24.8	1.00E-01	3.31E-03	4.85E-12	0
...							
zeroth_order	−1	−1	−1	9.98E-02	−1	−1	0

Notes. We only give a few representative rows out of a total of 45. The first row of column headings is the FITS table column names, whose units are in the second row. For compactness, we have omitted the error columns (err_count_rate, err_photon_flux, and err_energy_flux). A value of −1 indicates null data. Continuum band labels are a "c" followed by the band wavelength in mÅ.

capabilities of ISIS may be found in Houck & Denicola (2000) and Noble & Nowak (2008).

At the top level, there is a single ISIS function which determines the observational configuration and then calls the appropriate script to sequence the CIAO tasks. Customization can be done for non-standard extractions by setting global state variables (e.g., to specify coordinates other than given in the event file header, or to use an alternate source detection method, or to define alternate sky or extraction regions). Lower level functions can be easily scripted to provide nearly arbitrary CIAO tool invocations. While the *TGCat* processing software is not intended to be fully general or to replace the UNIX command-line CIAO interface, it can be used for many common processing scenarios. The processing software used to produce the *TGCat* archive is documented and made available to users[14] (see Section 4 for details).

2.2. Flux Properties Table

The flux properties table is a FITS binary table of count rates and fluxes integrated over bandpasses. Details of the file contents are given by a partial listing for one observation in Table 4. There is no continuum fitting or subtraction or line-profile fitting. Statistical uncertainties are provided for each band. For LETGS, counts are background subtracted, but flux values have not been computed since they are model dependent due to overlapping higher-order contributions in the counts. Regions are assigned a label indicating the dominant element and ion appropriate for a typical coronal thermal plasma in the region. Near each such band is a "continuum" band, labeled with a *c*. To obtain an approximate line-region rate or flux, the continuum values can be interpolated and scaled to the line bandpass using the tabulated wlo and whi values giving the bandpass, and then subtracted from the line rate. The band wavelength midpoint, $wmid$ is tabulated. (For continuum sources, the line labels are irrelevant.) Broadband rates have been given for the HEG, MEG, LETG/HRC, and LETG/ACIS spectral ranges. We provide zeroth-order rates, but it must be realized that for many sources observed with ACIS, the zeroth order is saturated by photon pileup (Davis 2001). Rates are supplied for the Chandra Source Catalog "science energy" bands (Evans et al. 2010) and are labeled with a $csc\_$ prefix and a band mnemonic character (*b*, *u*, *s*, *m*, and *h*).

If a bandpass is truncated by the sensitivity limits, or if the effective area is zero-valued somewhere within the band, then a value in the *flag* column will be non-zero (e.g., *letgs_band* for HETGS observations). These rates are intended to be used for qualitative characterization of spectra. Detailed modeling is required for determination of rigorous values, particularly for weak or heavily absorbed sources in the softer bands (>20 Å) where noise and background can be amplified in the flux estimate.

2.3. Archive and Process Management

Archive and process management are performed by an automated system that collects and processes observations after their proprietary period has expired, organizes information for manual verification and validation, and then responds to the manual review by either releasing to the public Web site, reprocessing with adjusted extraction parameters, or holding the output indefinitely. The major components include a MySQL database for storage of metadata, a file-based data archive, and the processing software described in Section 2.1. The data flow is sketched in Figure 1.

The information relevant to processing includes the extractions, source, and spectral properties tables, data product files (see Table 3), and queue tables. The extractions table has one entry per processed extraction. Any ObsID can have many sources and any source might be found in multiple ObsIDs. Any source in any ObsID could have multiple extractions (e.g., for time slices, or different cross-dispersion widths). The extractions table stores one entry for each combination thereof. In order to consolidate all extractions of a single source, there is a source table to associate sources by indexes on SIMBAD[15] identifier, a *TGCat* identifier, and source celestial coordinates.

An automatic process runs at regular intervals (typically daily), downloads a list of gratings observations from the public

---

[14] http://space.mit.edu/cxc/analysis/tgcat/

[15] http://simbad.u-strasbg.fr/simbad/

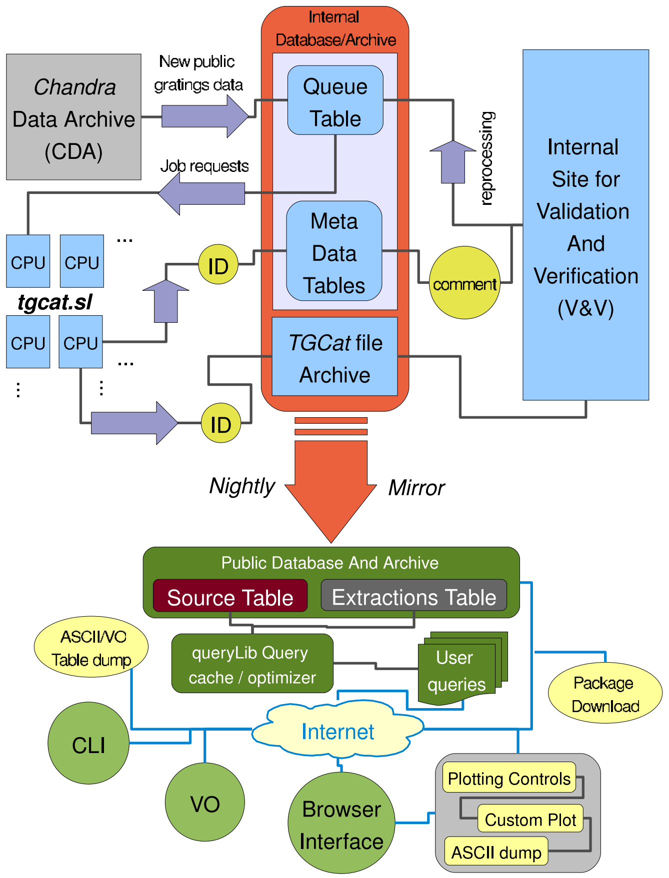


Figure 1. *TGCat* data flow. The upper portion of the diagram (above "Nightly Mirror") is the internal archive creation infrastructure. Below that is the public user interface. The *Chandra* Archive (upper left box) is an external component. It is the primary input to Internal Database/Archive control process (upper central rectangle), which then schedules the processing (middle-left array of CPU boxes), and populates the public database (lower portion of the diagram), after manual review by scientists (V&V). The Public Database and Archive provides different interfaces and options for viewing or downloading the spectra and associated information. No data are retrieved from the CDA until their proprietary period has expired.

Chandra archive, and compares it with the list of ObsIDs that have been submitted to *TGCat*. Any ObsID not in *TGCat* whose proprietary period has expired will be added to the queue for *TGCat* processing. Daemons (written in python) run on any number of the local network hosts and continuously request entries from the queue in a first-in, first-out manner. They then configure the queue entry (which can contain a number of custom processing parameters), setup work spaces and logging, fork off the processing to the ISIS pipeline scripts (described in Section 2.1), and then ingest products into the database and archive.

During the ingest step, several automatic checks are done to evaluate whether the processed data are complete. If so, metadata are added to the main extractions table which returns a unique identifier that is then used to tag the products and also link the source and extraction tables. For incomplete data, the queue entry is assigned an error and notification sent to *TGCat* administrators. *TGCat* operators have the choice to investigate the existing processing workspace or simply re-evaluate parameters and re-queue the extraction as new.

Verification and validation (V&V) is done by members of the *TGCat* science team (CXC scientists at MIT and SAO; HETG project scientists at MIT) in order to confirm the zeroth-order placement, proper sky mask and spectral binning regions, and to assess the presence of any confusing sources. Review is done through an internal administrative Web site similar to the public interface for browsing extractions, but with the addition of forms for changing extraction parameters and re-queueing processing (if necessary), and for marking the extraction with a "good," "bad," or "warning" label. The reviewer can edit comment fields which are available for internal adminstrative purposes only, and the ones which provide a V&V report for the end user (e.g., the "warning" category flags cases which cannot be mitigated by processing, but of which the user should be aware, such as double sources that cannot be separately extracted). The "bad" extractions are never shown on the public site; the administrators have the option of rejecting any extraction for any reason. Each extraction is tagged with the date of processing, the version of *TGCat* used for processing, and a group ID which is unique per group of extractions intended to be of the same object for a single ObsID, which allows us to keep track of accepted extractions and to avoid duplicate extractions. The V&V process uses the same summary graphical products that are presented to users in the catalog.

The processing parameters are saved for each extraction. This allows us to easily reprocess all or any subset of the archive in the event of software or calibration updates. Several such reprocessings have already occurred, motivated by a CALDB updates or CIAO software improvements. Reprocessing the entire archive takes about three days using one dedicated host and several others on the local network, as available.

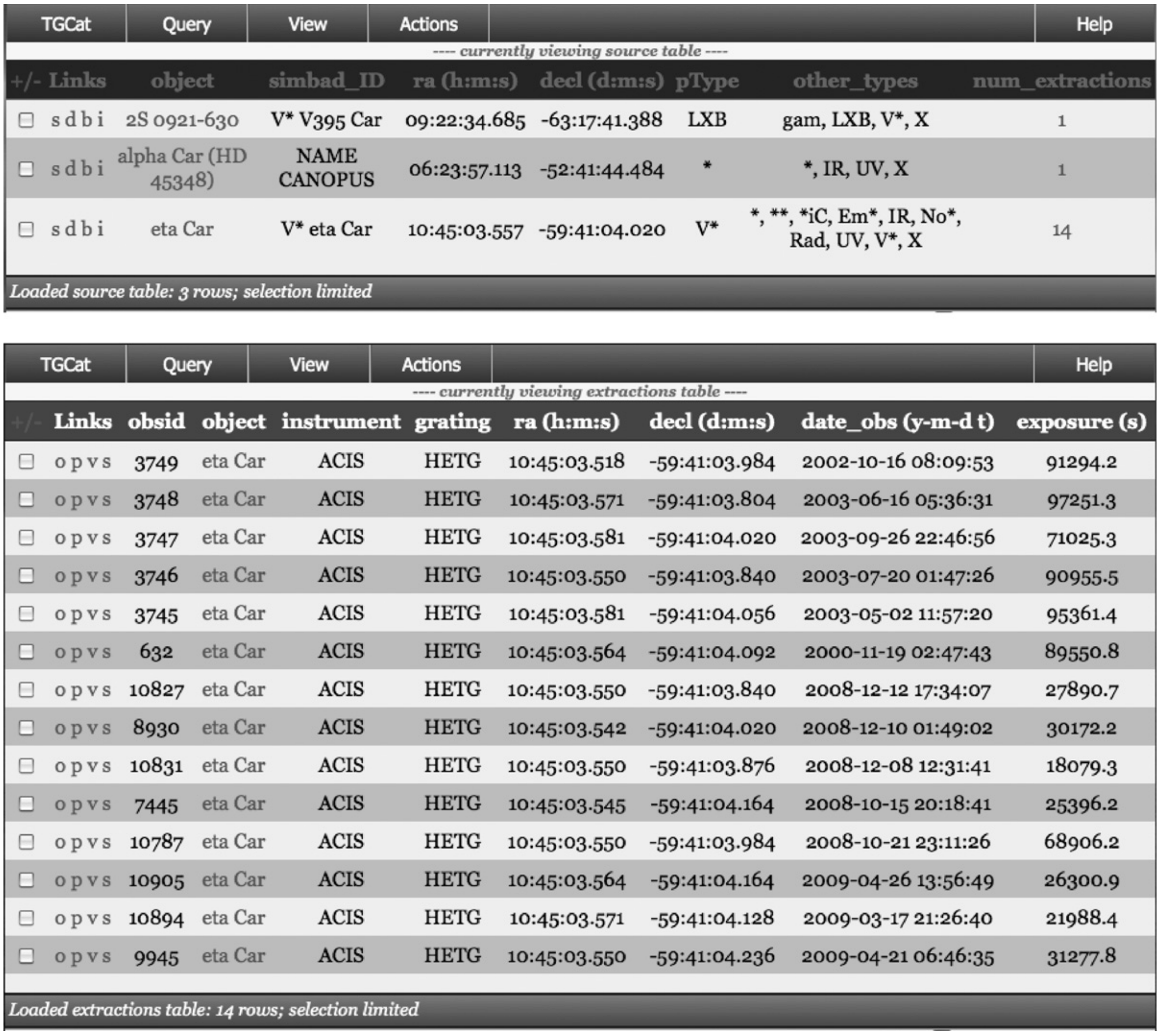
3. THE CATALOG

Once an extraction passes the V&V step it is ingested into the public archive and made accessible through the catalog interfaces. The web interface (http://tgcat.mit.edu/) allows searching on any of several different criteria, including a name string, cone search, object type (from SIMBAD), ObsID, any of a number of extraction or header quantities indexed in the database, or spectral properties (see Section 2.2). Boolean combinations of indexed quantities are also supported. For convenience, there is also a "Quick Search" form which allows some simple name substring, ObsID, or coordinate searches.

For successful object searches, a source table is returned with basic information, such as the name, coordinates, object types, external links (such as SIMBAD), and a link to the extractions (other types of searches, with different types of results are possible). Selecting the link on the "object" column opens the extractions table showing details of each extraction, which is typically one per ObsID. Here, we also have some relevant external links (e.g., to the observer's input observational parameters or to publications using the observation). We show examples of source and extractions tables in Figure 2.

Given an extractions table, there are several actions one can take. Selecting any single extraction opens a "Preview Gallery" of the pre-computed summary products. These are useful for assessing the quality of an observation via counts spectra or field images.

The preview plots include detail images of the zeroth-order region, field images in rotated sky coordinates, images in grating spectral coordinates (wavelength and cross-dispersion), light curves for source and background regions, counts spectra, and fluxed spectra. For LETGS, we provide one spectral histogram fluxed as if all counts were first order as a useful approximation of the flux, since the LETGS cannot resolve orders (the HRC-S detector has very low energy resolution). For ACIS observations, we also show an "order-sorting" plot, sometimes useful for



+/-	Links	object	simbad_ID	ra (h:m:s)	decl (d:m:s)	pType	other_types	num_extractions
☐	s d b i	2S 0921-630	V* V395 Car	09:22:34.685	-63:17:41.388	LXB	gam, LXB, V*, X	1
☐	s d b i	alpha Car (HD 45348)	NAME CANOPUS	06:23:57.113	-52:41:44.484	*	*, IR, UV, X	1
☐	s d b i	eta Car	V* eta Car	10:45:03.557	-59:41:04.020	V*	*, **, *iC, Em*, IR, No*, Rad, UV, V*, X	14

Loaded source table: 3 rows; selection limited

+/-	Links	obsid	object	instrument	grating	ra (h:m:s)	decl (d:m:s)	date_obs (y-m-d t)	exposure (s)
☐	o p v s	3749	eta Car	ACIS	HETG	10:45:03.518	-59:41:03.984	2002-10-16 08:09:53	91294.2
☐	o p v s	3748	eta Car	ACIS	HETG	10:45:03.571	-59:41:03.804	2003-06-16 05:36:31	97251.3
☐	o p v s	3747	eta Car	ACIS	HETG	10:45:03.581	-59:41:04.020	2003-09-26 22:46:56	71025.3
☐	o p v s	3746	eta Car	ACIS	HETG	10:45:03.550	-59:41:03.840	2003-07-20 01:47:26	90955.5
☐	o p v s	3745	eta Car	ACIS	HETG	10:45:03.581	-59:41:04.056	2003-05-02 11:57:20	95361.4
☐	o p v s	632	eta Car	ACIS	HETG	10:45:03.564	-59:41:04.092	2000-11-19 02:47:43	89550.8
☐	o p v s	10827	eta Car	ACIS	HETG	10:45:03.550	-59:41:03.840	2008-12-12 17:34:07	27890.7
☐	o p v s	8930	eta Car	ACIS	HETG	10:45:03.542	-59:41:04.020	2008-12-10 01:49:02	30172.2
☐	o p v s	10831	eta Car	ACIS	HETG	10:45:03.550	-59:41:03.876	2008-12-08 12:31:41	18079.3
☐	o p v s	7445	eta Car	ACIS	HETG	10:45:03.545	-59:41:04.164	2008-10-15 20:18:41	25396.2
☐	o p v s	10787	eta Car	ACIS	HETG	10:45:03.550	-59:41:03.984	2008-10-21 23:11:26	68906.2
☐	o p v s	10905	eta Car	ACIS	HETG	10:45:03.564	-59:41:04.164	2009-04-26 13:56:49	26300.9
☐	o p v s	10894	eta Car	ACIS	HETG	10:45:03.571	-59:41:04.128	2009-03-17 21:26:40	21988.4
☐	o p v s	9945	eta Car	ACIS	HETG	10:45:03.550	-59:41:04.236	2009-04-21 06:46:35	31277.8

Loaded extractions table: 14 rows; selection limited

Figure 2. Top: *TGCat* source table for "Quick Search" on the string, "car" showing three matching sources. The "Links" column provides easy access to external data, such as SIMBAD ("s"), NVO Datascope ("d"), XMM BiRD "b," or 2MASS and DSS images ("i"). Bottom: the extractions table obtained by selecting source "eta Car" from the upper table. Placing the mouse pointer over the object name (which links to the summary plots page) reveals a pop-up flux spectrum plot as a quick preview of the quality. Other "Links" take you to the *Chandra* observation catalog ("o"), publications ("p"), the V&V report ("v"), or SIMBAD ("s").

visualizing the intrinsic energy separation afforded by the CCD detectors. Figure 3 shows an example screen for the longest exposure η Car extraction shown in Figure 2. Each thumbnail image, when selected in the browser, expands into a larger view.

3.1. *Custom Plotting*

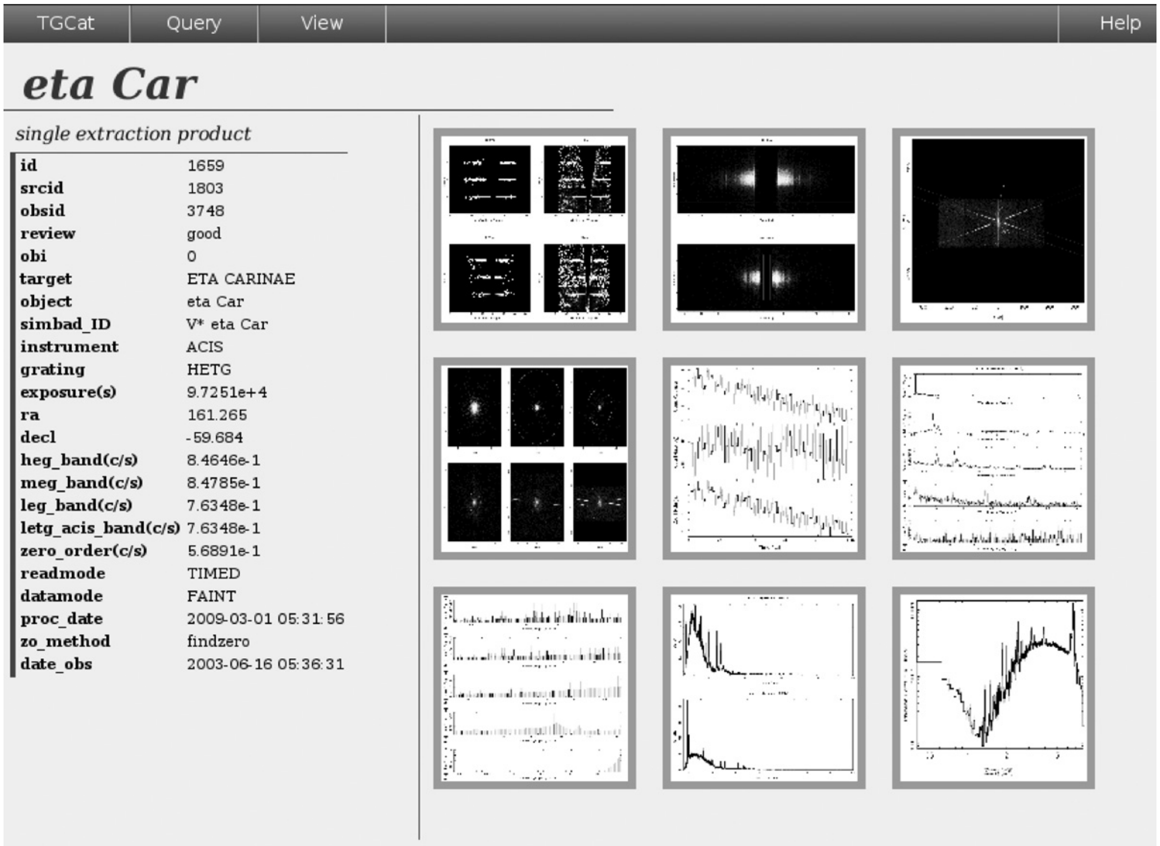
To go beyond simply assessing the data quality, we have provided an interactive and very flexible custom plotting interface. It can plot a single order, combined orders for a single observation, or it can combine spectra for multiple observations. Spectra can be binned by signal to noise and number of channels (or both). Scaling can be log or linear with user-defined ranges. There are diverse choices for spectral coordinates (frequency-like or wavelength-like) and flux units.

For careful and rigorous spectral analysis, we recommend download of the individual spectral and response data products (see Section 3.2). But for some applications, the interactive plot result itself may be sufficient, such as for inclusion in multi-band plots to show the X-ray spectrum beside radio, infrared, optical, or radio results. For such use, the displayed plot can be downloaded as a plain-text table with embedded comments which document the contents and units of the columns.

Behind the scenes, plotting is implemented on the server by taking the request for plotting parameters along with a unique file name, creating a small ISIS script of the commands for loading and plotting data, and piping this script into an ISIS process. We prevent malicious use by checking all parameter inputs for appropriate values, running the ISIS process as an unprivileged user, and checking the temporary file name for validity. Since the file name is known at the time of the request, the page simply needs to be reloaded to show the new plot. This web-generated ISIS plotting script is also available for inspection or download. It can be used in conjunction with any downloaded data products and *TGCat* software as a template for user re-use or customization. Figure 4 shows an example plot page.

3.2. *Data Download*

There are several options for data download. Generally, users will want to download more than one file at a time for detailed analysis, such as spectra and responses for a single observation, products for multiple observations of the same target, or a selection of several similar objects. To this end, *TGCat* runs a packager process that parses a queue table. Requested packages

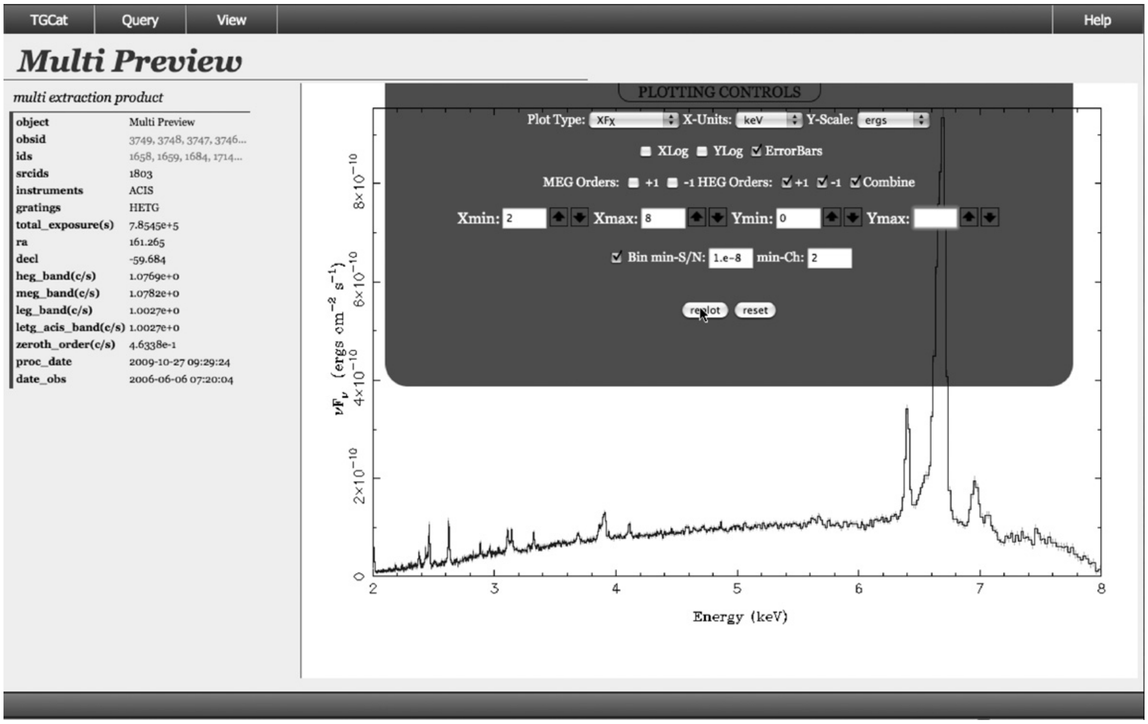


eta Car

single extraction product

id	1659
srcid	1803
obsid	3748
review	good
obi	0
target	ETA CARINAE
object	eta Car
simbad_ID	V* eta Car
instrument	ACIS
grating	HETG
exposure(s)	9.7251e+4
ra	161.265
decl	-59.684
heg_band(c/s)	8.4646e-1
meg_band(c/s)	8.4785e-1
leg_band(c/s)	7.63486-1
letg_acis_band(c/s)	7.6348e-1
zero_order(c/s)	5.6891e-1
readmode	TIMED
datamode	FAINT
proc_date	2009-03-01 05:31:56
zo_method	findzero
date_obs	2003-06-16 05:36:31

Figure 3. *TGCat* preview page for the longest η Car exposure. Each thumbnail is a link to a larger version. "Tooltips" appear to briefly describe each plot type when the mouse pointer hovers over a thumbnail, and a longer description of each plot type is available with the expanded view.



Multi Preview

multi extraction product

object	Multi Preview
obsid	3749, 3748, 3747, 3746...
ids	1658, 1659, 1684, 1714...
srcids	1803
instruments	ACIS
gratings	HETG
total_exposure(s)	7.8545e+5
ra	161.265
decl	-59.684
heg_band(c/s)	1.0769e+0
meg_band(c/s)	1.0782e+0
leg_band(c/s)	1.0027e+0
letg_acis_band(c/s)	1.0027e+0
zeroth_order(c/s)	4.6338e-1
proc_date	2009-10-27 09:29:24
date_obs	2006-06-06 07:20:04

Plot Type: XFx X-Units: keV Y-Scale: ergs

☐ XLog ☐ YLog ☑ ErrorBars

MEG Orders: ☐ +1 ☐ -1 HEG Orders: ☑ +1 ☑ -1 ☑ Combine

Xmin: 2 Xmax: 8 Ymin: 0 Ymax:

☑ Bin min-S/N: 1.e-8 min-Ch: 2

Figure 4. Interactive spectral plot for the combined HEG positive and negative orders for the 14 observations of η Car listed in the extractions table shown in Figure 2. The "PLOTTING CONTROLS" form is shown in the active state for selection of scales, ranges, units, and binning parameters. The displayed data can be captured as a PNG-file, dumped to an ASCII table, or the ISIS plot script can be retrieved.

are added to the package queue table, validated and read by the packager which then fetches data to a temporary space, placing them in a directory hierarchy tagged with ObsID and *TGCat* ID. The entire hierarchy is then tarred, compressed, and provided to the user along with file checksums via HTTP. Users have the opportunity to select the types of files to package, such as only spectra and responses, or to also include event lists. It is also possible to view and download individual files from a file-table list.

3.3. *Virtual Observatory Support*

TGCat is a registered Virtual Observatory (VO) service providing both the "Simple Cone Search" and "Simple Image Access" protocols. These are both implemented as XML output plug-ins taking input from a PHP script that provides appropriate parameters, then creates a query object after parsing the input data, and finally runs the query in exactly the same way as for the web and ASCII interfaces. Error handling is done by the calling script. Query types are tracked so we can obtain statistics on requests coming from services such as the HEASARC Datascope[16] VO search facility. The VO interface further allows searching and data retrieval through any VO-enabled program.

4. USER SOFTWARE

All the software used to drive the CIAO reprocessing and to generate summary products is available for download.[17] This also includes some utility scripts for retrieval of data from the CDA and for setting up the directory for processing. The software makes use of S-Lang and some S-Lang modules (such as the `curl` module for data retrieval), and of ISIS for controlling the CIAO processes and for handling spectra to make plots. A user may wish to run their own extractions if, for example, time filtering of variable sources is required, or if extracting serendipitous sources. Users can also apply the *TGCat* scripts to their own proprietary data not yet available from *TGCat*. Once software has been installed, a basic reprocessing session is quite simple; we provide an example in Appendix A.

5. EXAMPLE APPLICATIONS OF *TGCat*

5.1. *Real Science of Hot Stars*

Walborn et al. (2009) used *TGCat* to compare X-ray spectra of O-stars with respect to spectral type and luminosity class. For each star in their comparison, they added all available spectra of each object to improve the signal. Using spectra available in *TGCat* ensured they were processed identically with the most appropriate calibration for each separate observation.

This technique has recently been used by J. S. Nichols et al. (2011, in preparation) to add the seven spectra of HD 93129A, an O2 I star. The addition of these seven separate exposures allowed modeling of the He-like and H-like emission lines in spite of the low signal to noise of the individual spectra. These authors found that the emission lines can be modeled as skewed, apparently shifted lines consistent with a dense, absorbed wind, much like that seen in the O4 I star, ζ Pup. W. L. Waldron et al. (2011, in preparation) have used spectral data for the B-type supergiant, κ Ori for a detailed analysis of the wind. This method of modeling combined spectra is a standard feature of ISIS.

TGCat is particularly useful for variability analysis of emission line regions. Assuming sufficient signal to noise, the researcher can filter an observation into approximately equal exposure times. Each time interval can then be processed with *TGCat* scripts to apply appropriate calibrations and to extract the spectra. Variability tools can then be used on the collection to look for spectral variability.

Table 5
TGCat Archive Statistics

Quantity	Value
Date	2010 Jul 1
Total ObsIDs	1063
Distinct sources	328
ACIS/HETG observations	738
ACIS/LETG observations	98
HRC/LETG observations	227
Minimum exposure	52 s
Mean exposure	45 ks
Maximum exposure	171 ks
Number of archived files	36,397
Archive size	158.95 GB
Space per ObsID	153.12 MB

The fields of view of the entire set of *Chandra* HETG observations contain zeroth-order and dispersed spectra of a large number of non-targeted stars. While *TGCat* does not yet contain systematic extractions of these spectra, there are plans to provide them. These "bonus" spectra represent a large untapped resource of the *Chandra* project. One example of this type of project is the collection of HETG spectra of θ^2 Ori A from multiple HETG pointings on the Orion Nebular Cluster. When all of the θ^2 Ori A spectra that fell in any of these HETG field of view were extracted, the cumulative exposure was 520 ks, resulting in a high very high quality spectrum (Mitschang et al. 2011).

5.2. *Anomalous Adara*

Early in the development and testing of *TGCat*, one of us (H.L.M.) did a selection on LETG/HRC-S observations and sorted by descending exposure time. At the top of the list was ϵ CMa (Adara, ObsID 6441). Browsing the images, we noticed that there was diffuse emission surrounding the spectrum. This was unexpected for a B-star—a point source. Further investigation of the *TGCat* archive found no other sources with extended emission near 30 Å, nor could we identify any instrumental effect particular to this observation. This led to a successful ACIS-S imaging spectroscopy proposal (ObsID 9926) to search for the spectral excess near 0.4 keV and any spatially resolved emission (since the ACIS PSF is simpler and better modeled than that of LETG zeroth order). The result was negative—the spatial/spectral feature was not physical; while not understood, it certainly was not due to terribly interesting and unprecedented astrophysics of Adara. We put the study aside as an unreproducible artifact. Some time later, another bright B-star was observed, and the PI contacted us (L. Oskinova 2010, private communication) to ask what could be the bump in the LETG/HRC-S spectrum of α Cru (ObsID 8937)? This rang a bell. We finally came up with a plausible instrumental effect which would only affect UV-bright B-stars (very high order diffraction by the LETG coarse-support structure), and the Calibration Group provided some supporting quantitative analysis (Drake 2010). Case closed.

6. CONCLUSIONS

As of this writing, the *TGCat* archive contains extractions from over 1000 observations. Figure 5 shows a sky map of

---

[16] http://heasarc.gsfc.nasa.gov/cgi-bin/vo/datascope/init.pl

[17] http://space.mit.edu/cxc/analysis/tgcat/

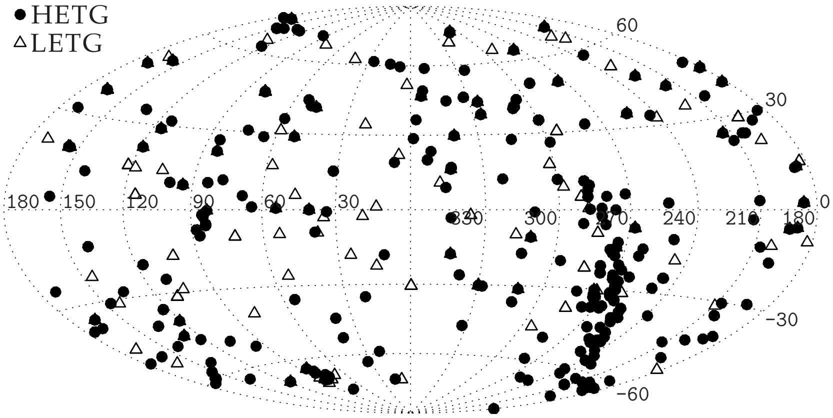


Figure 5. Sky map of *Chandra* grating observations included in *TGCat*, shown in a Hammer projection of celestial coordinates, right ascension (x-axis), and declination (y-axis). The concentration around $(\alpha, \delta) = (270, -30)$ stands out; it is the direction toward the Galactic center.

Table 6
Summary Observation Catalog

Column Number	Description
1	*Chandra* observation identifier
2	*TGCat* preferred object identifier
3	SIMBAD primary object identifier
4	SIMBAD object primary type
5	*Chandra* instrument designation
6	*Chandra* grating designation
7	Right ascension (J2000)
8	Declination (J2000)
9–14	Observation start date
15	Exposure time
16	Zeroth-order centroid position determination method
17	MEG-band count rate (1.7–25 Å)
18	MEG-band count rate error
19	MEG-band energy flux (1.7–25 Å)
20	MEG-band energy flux error
21	HEG-band count rate (1.7–15 Å)
22	HEG-band count rate error
23	HEG-band energy flux (1.7–15 Å)
24	HEG-band energy flux error
25	CSC b-band count rate (2.0–24.8 Å)
26	CSC b-band count rate error
27	CSC *b*-band energy flux (2.0–24.8 Å)
28	CSC *b*-band energy flux error
29	CSC *s*-band count rate (10.33–24.8 Å)
30	CSC *s*-band count rate error
31	CSC *s*-band energy flux (10.33–24.8 Å)
32	CSC *s*-band energy flux error
33	CSC *m*-band count rate (6.2–10.33 Å)
34	CSC *m*-band count rate error
35	CSC *m*-band energy flux (6.2–10.33 Å)
36	CSC *m*-band energy flux error
37	CSC *h*-band count rate (2.0–6.2 Å)
38	CSC *h*-band count rate error
39	CSC *h*-band energy flux (2.0–6.2 Å)
40	CSC *h*-band energy flux error

Notes. Shown here is a description of the columns in the *TGCat* observations catalog.

(This table is available in its entirety in machine-readable and Virtual Observatory (VO) forms in the online journal. A portion is shown here for guidance regarding its form and content.)

Chandra grating observations included in *TGCat* and Table 5 gives a few statistics on the archive. In Table 6, we give a listing of the observations included to date, along with basic source information and count rates collated from the flux properties files. These rates and fluxes are for combined orders, and for HETGS, combined HEG and MEG spectra.

TGCat represents a valuable resource for analysis-ready products for experienced X-ray high-resolution spectroscopists as well as a place for others to browse, explore, and find information relevant to research in other spectral regimes.

While *TGCat* is a relatively mature resource, it is not finished. It will, of course, grow as new observations become public. But we also have plans for enhancements. There are a good number of serendipitous sources to extract, as well as multiple sources in crowded fields, and some multiple observations for which combined products would be useful. We also intend to provide low-resolution spectral extractions for zeroth orders.

Support for this work was provided by the National Aeronautics and Space Administration through the Smithsonian Astrophysical Observatory contract SV3-73016 to MIT for Support of the Chandra X-Ray Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the National Aeronautics Space Administration under contract NAS8-03060. J.S.N., A.M., and D.M. acknowledge support from the *Chandra* X-ray Center NASA contract NAS8-03060. Support for the Adara work was provided by the National Aeronautics and Space Administration through *Chandra* Award Number GO9-0012X issued by the *Chandra X-ray Observatory* Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the National Aeronautics Space Administration under contract NAS8-03060.

Facility: CXO (HETGS, LETGS)

APPENDIX

EXAMPLE PROCESSING SCRIPT

Consider obtaining two observations with ObsID 5 and 1103. Assume that for ObsID 5, all defaults are appropriate, but that for ObsID 1103, we need “`findzo`” because zeroth order is distorted by pileup. As the pipeline processes are run, the CIAO commands are echoed to the terminal, and can be captured for further customization, if desired. The steps are as follows.

```
unix> # retrieve the data from cdaftp and
configure directories:
unix> download_obs 5 1103
```

```
unix> # start isis, load software:
unix> isis
isis> require("tgcat");

isis> % run standard extraction for ObsID 5:
isis> run_pipe("obs_5");

isis> % set alternate detection method and run
extraction:
isis> s=set_source_detection_info("findzo");
isis> run_pipe("obs_1103"; detect_info = s);

isis> % spawn a look at summary products (using
the external "ImageMagick" package):
isis> ! display obs_5/summary*.ps
isis> ! display obs_1103/summary*.ps

isis> % continue with detailed analysis:
isis> d = load_set_acis("obs_5", [3,4,9,10]);
% rows 3,4,9, &10 are HEG, MEG 1st orders.
isis> list_data;        % list the loaded
spectra
isis> plot_data(1);  % plot the HEG -1 order
counts spectrum
isis> require("xspec");
isis> fit_fun("(gauss(1)+poly(1)) * wabs(1)");
isis> % etc ...
```

REFERENCES

Brinkman, A. C., et al. 2000, ApJ, 530, L111

Canizares, C. R., et al. 2005, PASP, 117, 1144

Davis, J. E. 2001, ApJ, 562, 575

Drake, J. 2010, Chandra Newslett., 17, 21

Evans, I. N., et al. 2010, ApJS, 189, 37

Fruscione, A., et al. 2006, Proc. SPIE, 6270, 62701V-1

Hanisch, R. J., Farris, A., Greisen, E. W., Pence, W. D., Schlesinger, B. M., Teuben, P. J., Thompson, R. W., & Warnock, A., III. 2001, A&A, 376, 359

Houck, J. C., & Denicola, L. A. 2000, in ASP Conf. Proc. 216, Astronomical Data Analysis Software and Systems IX, ed. N. Manset, C. Veillet, & D. Crabtree (San Francisco, CA: ASP), 591

Huenemoerder, D. P., Mitschang, A., Nichols, J., Nowak, M. A., Schulz, N. S., & Dewey, D. 2009, Chandra Newslett., 16, 33

Mitschang, A. W., Huenemoerder, D. P., & Nichols, J. S. 2010, arXiv:1001.0039

Mitschang, A. W., Schulz, N. S., Huenemoerder, D. P., Nichols, J. S., & Testa, P. 2011, ApJ, in press

Nichols, J. S., Huenemoerder, D., & Ishibashi, K. 2008, in AAS/High Energy Astrophysics Division, Vol. 10, 13.15

Noble, M. S., & Nowak, M. A. 2008, PASP, 120, 821

Paerels, F. B. S., & Kahn, S. M. 2003, ARA&A, 41, 291

Walborn, N. R., Nichols, J. S., & Waldron, W. L. 2009, ApJ, 703, 633

Weisskopf, M. C., Brinkman, B., Canizares, C., Garmire, G., Murray, S., & Van Speybroeck, L. P. 2002, PASP, 114, 1

Westbrook, O. W., et al. 2008, ApJS, 176, 218